Modelling Stress Recognition in Conflict Resolution Scenarios

Marco Gomes, Davide Carneiro, Paulo Novais and José Neves

Department of Informatics, University of Minho
pg18373@alunos.uminho.pt, {dcarneiro,pjon,jmn}@di.uminho.pt

Abstract. The current trend in Online Dispute Resolution focuses mostly on the development of technological tools that allow parties to solve conflicts through telecommunication means. However, this tendency leaves aside key issues, namely the context information that was previously available in traditional Alternative Dispute Resolution processes. The main weakness of this approach is that conflict resolution may become focused solely on objective issues. In order to overcome this inconvenience, we move forward to incorporate context and behavioural information in an Online Dispute Resolution platform. In particular, we consider the estimation of the level of stress and the prediction of the stress state evolution. As a result, the conflict resolution platform or the mediator may predict to what extent a party is affected by a particular matter, allowing one to adapt the conflict resolution strategy to a specific scenario in real time.

Keywords: Hybrid Artificial Intelligence Systems, Online Dispute Resolution, Stress, Cognitive Activation Theory of Stress

1 Introduction

Online Dispute Resolution (ODR) is a form of dispute resolution that takes place partially of wholly in a digital environment [**?**]. The use of technological solutions in this field is nowadays well established. However, the current trend continues to focus mainly on the development of technological tools. As a result, the actual ODR systems leave aside important issues that are present in traditional dispute resolution processes, namely, the context-dependency issue. This issue has a preponderant role in human behaviour. The omission of context and behavioural information can influence the course of action and, consequently, the outcome of a conflict resolution scenario. The use of a synergistic combination of multiple Artificial Intelligence techniques [**?**] [**?**] and, more particularly Ambient Intelligence techniques, can help to suppress this lack.

In order to address this challenge this work aims to develop mechanisms that operate in the virtual environment of ODR to collect context information and perceive the activities being performed. Basically, the underlying intent is to extend the traditional technology-based conflict resolution, in which a user simply interacts with the system, with a new component: an intelligent environment. These environments are pervasive and transparent, i.e. a person should not perceive the environment in any other way than by the actions it executes. This is important since when people are aware that they are under monitorization, they tend to behave differently.

With this approach it is expected to provide useful context knowledge to support the lifecycle of the conflict resolution model. The main objective is to capture context information that can be used by conflict resolution platforms to achieve better and more satisfactory outcomes for the involved parties [**?**]. This will include the estimation of the level of stress, the attention level, behavioural patterns, among others. Being able to make assumptions about the information that can be used to characterize a person's situation is extremely relevant to assist a conflict resolution platform finding the path to achieve successful outcomes [**?**]. From the point of view of the (electronic or human) mediator, the access to this information is also vital in order to plan the right strategy and perceive how each issue affects each party, much like it is done in traditional conflict resolution processes, when parties meet face-to-face.

2 The Stress Interpretation and a Multimodal Approach

Many experts endorse the original definition of stress concept to the proposed by Hans Selye [**?**], who coined the term as it is presently used. He defined the stress like a non-specific response of the body to any demand placed upon it. Selye conceived of external demands as *stressors* (the load or stimulus that triggered a response) and the internal body changes that they produced as the *stress response* (triggered by a load or a stimulus). He was the first person to document the chemical and hormonal changes that occur with stress. However the definitions provided were not conclusive for the entire scientific community. The concept of *stress* is still an open discussion in the scientific community. The main reason for this lies in the multiplicity of factors and the subjective nature of stress phenomenon, which led to multiple interpretations. With so many factors that can contribute to stress it can be difficult to define it. In this open discussion some argue that the *stress* concept is elusive because it is poorly defined [**?**] and others prefer to not define the concept until stress research reach a consensual significance. Meanwhile, scientists have circumvented the problem of a clear and agreed definition of stress by defining it empirically. It's difficult to measure stress if there is no agreement on what the definition of stress should be, but this does not prevent the advancement of science in this area but this has not stopped the advance in this field. Researches start to focus upon cognitive and behavioural causes for stress, and stress became viewed as a mind-body, *psychosomatic*, or psycho-physiologic phenomenon. A free interpretation of this phenomenon could refer stress as a physico-physiologic arousal response occurring in the body as result of stimuli, and these stimuli become a *stressor* by virtue of the cognitive interpretation of the individual. This will be our interpretation of stress that will serve as a cientific starting point to modelling stress in the conflict resolution scenarios.

Some experimental results [**?**] demonstrate that single-modality features are not sufficiently precise for stress recognition, while the integration of multiple-modality features are very helpful for accurate stress recognition. The following generic diagram (Fig. 1) represents a multimodal approach to the stress recognition problem. It consists of two portions, the leftmost portion, from left to the "stress" node, depicts the elements that can alter human stress. These elements are included in "context" node and represent the generalization of two main categories (sources of contextual information): the "user-centric" context and "environmental" context. On the other hand, the rightmost portion of the diagram, from the "stress" node to the right portion, depicts the observable features that reveal stress. These features include quantifiable measurements on the user's physical appearance, physiology, behaviours and performance.

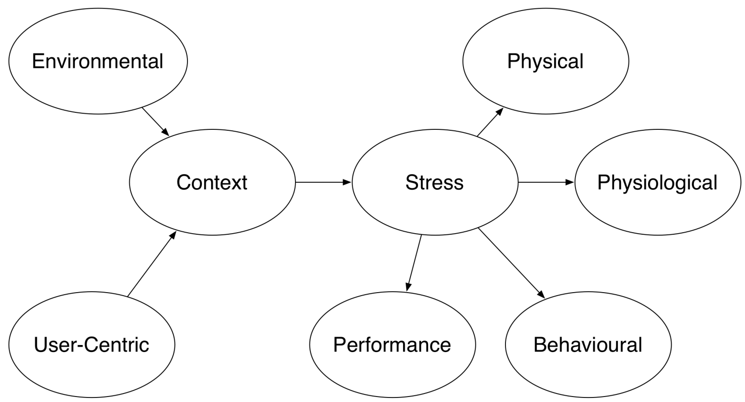

Fig. 1. A generic diagram for representing multimodality space in the recognition stress model. Due the space limit we don't draw all the features considered

The "context" node is divided in two types according to the source of contextual information, namely, the "user-centric" and the "environmental" context. User-centric information is composed of two categories: the background and the dynamic behaviour. The background is composed by several attributes that can be extracted from the user's profile. These attributes are the age, gender, working area, social status, personality traits among others. The dynamic behaviour reflects the contextual attributes related to the user's activity within the conflict resolution platform. These can be depicted by the intention of the user (if he/she really wants to achieve an outcome) and by the activity (if the user is an active or passive party in the process). The "environmental" information fuses the physical environment characteristics, social environment information and computational environment measurements. Physical environment includes attributes such as the time, temperature, location, noise level, and luminance. High levels of noise, extreme temperatures and low levels of luminance are well known potential stressors. The social environment is related to the social conditions when a user interacts with the system, like the population density when a user interacts with the platform (number of surrounding people per unit of area). The computation environmental context can be characterized by the measurement of the electromagnetic field and the number of surrounding electronic devices.

Among the features that can reveal stress, those that can characterize the "behavioural" node are the user's conflict style [**?**] (a conflict style can be a coping strategy in response to stressful conflict [**?**]), his/her interactions with the computer, the mouse/touch screen pressure from clicks/touches, his/her agitation level (through the sensory data from the accelerometer placed in mobile devices or by analyzing movement), as well as input frequency and speed. Also, the "performance node" is depicted in terms of accuracy and response, where the accuracy feature is related to the amount of touches in particular areas in the platform interface and the response feature corresponds to the analysis (qualitative and temporally) of user's responses to the conflict resolution demands. The physiological variables provide observable features about the user's stress state [**?**]. These features can be Galvanic Skin Response (GSR), that assesses the electrical proprieties of the skin in response to different kinds of stimuli, and General Somatic Activity (GSA) that assesses the minute movement of human body and many others such as respiration, pupilographic activity among others. Physical appearance includes the visual features that characterize the user's eyelid movement such as pupil movement (e.g. eye gaze, papillary response), facial expression or head movement. In addition, this is flexible enough to allow the insertion and modification of variables. For example, the variables under the behavioural node may vary with different applications, and the variables under the physiological node may change, depending on the availability of the requiring measuring devices.

3 Dynamic Stress Recognition Model

The ability to recognize human stress can have a determinant role in conflict resolution process. Having the information about the level of stress of the disputant parties is quite important in order for a mediator to correctly understand how each issue or event is affecting each party.

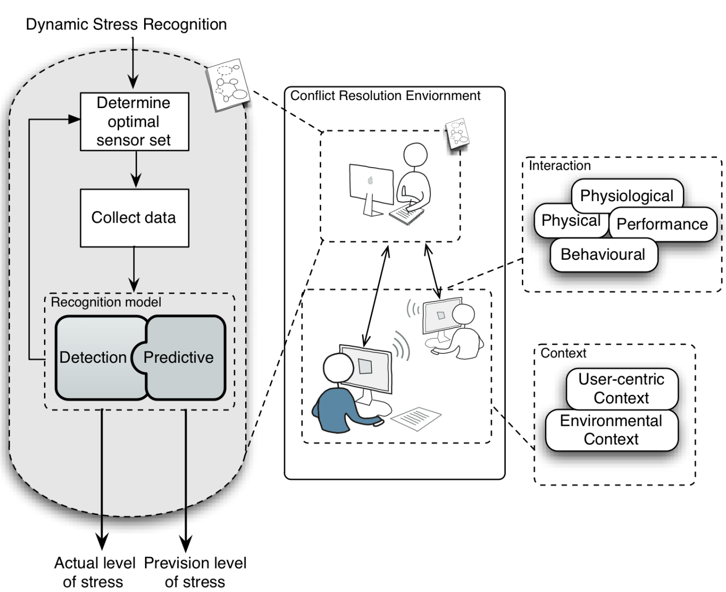

Fig. 2. An overview of the dynamic stress recognition system embedded in an intelligent environment.

Various approaches have been developed to recognize and detect user stress [**?**] [**?**] [**?**] [**?**]. The approaches or systems differ from each other in either the evidence modalities, or detection techniques, or both. Overall, our approach differs from the cited ones in that it employs the dynamic techniques to unify detection with stress prediction, utilizes evidences from multiple modalities, and is validated with psychology theories.

In order to enrich the conflict resolution process with this ability, one must consider the detection and prediction of stress evolution in real-time. These two complementary modules compose the stress recognition model. The first one, tries to measure in each time instance the stress, and the second gives an outlook as stress is evolving, based on the Cognitive Activation Theory of Stress (CATS). Based on this information, the (electronic or human) mediator is able to perceive how the state of each party is evolving. Figure 2 streamlines the mains procedures in applying the dynamic stress recognition to conflict resolution environment. At each time t, the platform performs three procedures - sensor optimal set, stress recognition, and returns information. Specifically, the system decides an optimal sensory action set to collect data with a sensor selection strategy. The collect data are propagated through the stress recognition model. Through this model, the stress detection module computes the stress value at time t and the stress prediction model (based in CATS) determines the tendency of stress evolution for time $t+1$. After the provision of information to the users, their stress levels may change and new evidences need to be collected. Thus the system goes to the next step and repeats the three procedures.

3.1 The Stress Detection Module

Sensor measurements inherently incorporate varying degrees of uncertainty and are, occasionally, spurious and incorrect. It is extremely important to select just the necessary set of sensors that can provide the more accurate and consistent data in each moment. To fulfil this restriction, the system must select, in each phase of the stress recognition process, a subset of sensors that is the most relevant and cost effective. Determining the most informative and cost effective subset of sensors requires an evaluation of all possible subsets of sensors, which is computationally intractable. To address the problem of sensor selection on the application layer we base our approach in Yongmian Zang and Quiang Ji [**?**] sensor selection methodology. Specifically, the authors propose four steps: (1) a Bayesian Network to represent the sensor and their dependencies; (2) a statistical measure to quantify the pairwise sensor synergy; (3) the construction of a synergy graph; and (4) a greedy algorithm is applied to identify the near-optimal subset of sensors. Some experiments were performed [**?**] and demonstrated that the least upper bound can approximate closely of the mutual information value. This permits to circumvented the computationally intractable problem of assess the mutual information value of all variables. The presented sensor selection approach reduces the computation time with minimum loss accuracy, comparing with others methods like the random methods and the brute force approach. In the other hand, this methodology, as presented for Yongmian Zang and Quiang Ji [**?**], have some limitations. Namely, the assumption of all sensors have the same cost benefice and the non-existence of any conditionally dependence between them. These limitation can be considered strong barriers to implementing a real-time conflict resolution system. Thus, some adjustments will exist in the future work.

Processing of Sensor Data. The management of the raw data captured by the selected sensors is a challenge in stress recognition, as well in many other scenarios. The differences among evidences (observable features that are capable of providing clues about the person's stress state) and within the same evidence can be from temporally and spatially separated experiences. To overcome this, the processing of sensor data must be well managed in order to make the information and accuracy losses minimum in the process of transforming raw data. In ODR systems events can occur within very different temporal windows (e.g. a party may wait for a proposal a few minutes or a few hours or days). In that sense we defined dynamic temporal windows. Instead of predetermining a static value for the duration of the windows, we support that the temporal windows must be adaptive, i.e this approach is event-oriented. It is clear that independently of the size of the temporal window, the system continues to select and record the sensor measurements. The difference resides in the times in which the other stages of the stress recognition system are fed. In our view, the most appropriate (considering the ODR characteristics) time is the interval between events that occur in the platform. In the remaining of the document, the letter t will denote the previously referred temporal window.

Feature Extraction. For each feature, a different extraction technique can be used. However, some techniques can be applied to more than one feature. For example, when there is a loss of data an interpolation method (curve fitting or regression analysis) could be used to fill the gap. Due to the specific sensory data that derives into multiple and different values, we just highlight some of the features considered. The values of physiological features like the GSR and the Pupillary Activity (PA) can be obtained by the analysis of the mean value of the GSR and PA measurements. Visual feature extraction starts with eye detection and tracking, which serves as the basis for subsequent eyelid movement monitoring, gaze determination, and face orientation estimation. Behavioural features can be elicit through the analysis of the negotiation activity and the parties' interactions with the platform. Context features can be extracted from the party's profile using a simple form to capture the age, gender, working area, social status, personality traits, among others. After this process, all the features are normalized to a value in the range [0, 1] using max-min normalization. Not all the values of features need to update further stages of the stress recognition model. This is due to the fact that these values may not vary much between time $t-1$ and t and, also, because some features assume static values. For example, the contextual information has some attributes that will not vary during the session in the conflict resolution platform, like the gender, the age, the work occupation among others. Thus being, the extraction techniques will only act after the verification of a significant variation in these parameters. This verification must be performed by an Analysis of Variance technique (e.g. ANOVA) to performs the role of selecting, with a certain threshold, which values of features should be taken into account in computing the variable value.

The Figure 3 gives an overall vision of how sensory data is processed, from the near-optimal sensor selection to the analysis of variance. The main function is to provide normalized data to compute the stress value. In the next section, the variables (modalities) will be the normalized values of all features analysed in this process.

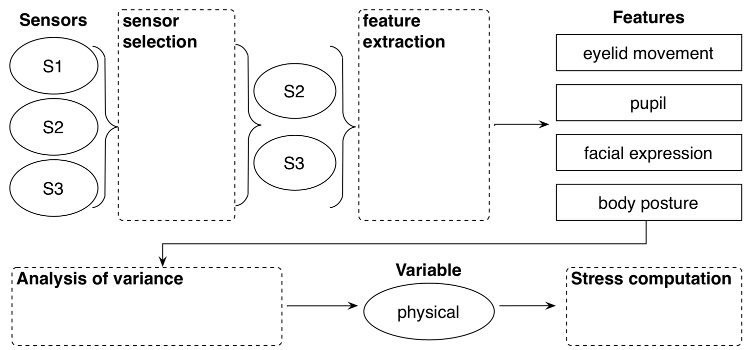

Fig. 3. The conceptual elements of the stress computation process

Decision Level Integration. To integrate all the multimodal evidences referenced in this work (see Sect. 2) we have chosen a decision level integration strategy. This strategy allows asynchronous processing of the available modalities, provides greater flexibility in modelling, and allows adaptive channel weighting between different modalities based on environmental conditions [**?**]. Our strategy uses a weighted sum to estimate the stress value with a temporal approach.

$$stress_t = \sum_{i=1}^{t-1} \left[c(i)_{wc} + b(i)_{wb} + p(i)_{wp} + ph(i)_{wph} + phy(i)_{wphy} \right] \tag{1}$$

In (1) we use the first characters of the modality (variable) name to refer the component that represents the sum of the normalized values from all the extracted features, i.e. c referring to context, p referring to performance, and so on. In addition, we subscript a variable by a step t to refer to the variable at time t. The w's, followed with the letters that represents the variables, denotes the weight of each component. At the beginning, the weights are equally distributed: every variable's weight is 0.2. Meanwhile, in the scientific literature the influence of each modality in the estimation of stress is not specified and has not been adequately explored yet. But this is not an irrelevant factor, in terms of accuracy of stress recognition systems. The relevance comes from the strong relationship between the interpretation of an individual's stress with the person and situation dependencies issues. To suppress this lack, we propose a novel approach to calculate the degree of influence by using Artificial Neural Networks (ANN). For each modality, an ANN (trained with synthetic datasets) is used in order to calculate the values of the weights. The input data includes all the values of all variables in time

$t-1$ and the output must be an estimation of the degree of influence of each variable. This value must be within the range [0,1] and the sum of all estimations values at time t must be 1. The resulting values represent the rate variation for each variable used in the computation of the stress value at $t-1$. Indeed, this approach tries to give a dynamic performance in the calculation of the weights. Thus, we fulfil the requirement of respecting the person and case dependencies by suppressing, heuristically, the lack of ground-truth in this issue. However, optimization and validation are needed. So, we aren't closed to other options.

3.2 A Stress Prediction Module Based in CATS

The Cognitive Activation Theory of Stress (CATS) presents a formal set of definitions that tries to reduce the group of terms, which may cover the same stress phenomenon. Therefore, CATS offers meanings formulated in symbolic logic, which is a natural advantage (from a computational point of view) in an area so subjective as the stress phenomenon. This is *cognitive* stress theory because CATS assumes that the stress response depends on acquired expectancies of the outcomes of stimuli and available responses, which can be regarded as acquired (learned) relations between stimuli, and between responses and stimuli. It is an activation theory since it is based on neurophysiological activation and arousal concepts.

According to CATS, the stress response (alarm) depends on the individuals appraisal (evaluation) of the stress situation. Formally, CATS stipulates that the alarm occurs when there is a discrepancy (D in Eq. 2) between what is expected or the *normal* situation (set value) and what is happening in reality (actual value). The *normal* situation is understood as the classification obtained by collecting reference values from the tests performed by normal persons in a normal (non-pathological) state. Symbolically, it is the difference (Eq. 2) between the value a variable should have (set value SV), and the real value (actual value AV) of the same variable.

$$D = (SV - AV); \tag{2}$$

In our approach the predicted value at time $t+1$ is the discrepancy (D) value between the stress actual value (AV) obtained at time t and the CATS set value (SV) at $t+1$, during the events (e.g. send a proposal, receive a proposal, reject, accept, etc.) sequence on the resolution conflict lifecycle. The actual value of stress (AV) is calculated by the stress detection module (see Sect. 3.1) and the set value (SV) by Eq. 5. The level of alarm depends on expectancy (Eq. 3) of the outcome of stimuli and the specific responses available for coping. When the subject has learned (stored) that one stimulus ($S1$) predicts the occurrence of another stimulus ($S2$) this is referred to as stimulus expectancy. The stimulus expectancy (3) is expressed as ${}_{S1}E_{S2}$ and is equivalent to the stimulus expectancy $S1-S2$, which means that $S1$ implies $S2$.

$${}_{S1}E_{S2} = (S1 \rightarrow S2); \tag{3}$$

$$H({}_{S1}E_{S2}) = SL_{S1} +_{S1} CT_{S2} +_{S1} PP_{S2}. \tag{4}$$

In our approach expectancies are quantified by several dimensions: acquisition strength, saliency of the event, and perceived probability. The acquisition strength of an expectancy (H) expresses that expectancies are acquired, according to the general principles of learning theory. This is calculated through the saliency of the event (SL), the contiguity of the event presentations (CT) and how often the events are occurring together (perceived probability PP). Primarily, the salience (SL) of an event can be achieved through an attention analysis, i.e., gathering the level of attention/interest when a party is experiencing the event. We can get this by using some elicited visual features like the gaze direction, the degree of eye open, and the size of pupil relative to luminance to classify the attention/interest level that arises from the event. The contiguity of the events represents the distance (temporal or spatially) between two events. Through a machine learning technique, and using a dataset extracted of our ODR prototype, we are able to distinguish between a contiguous and non-contiguous event. The perceived probability expresses the probability of the expected event, as it is perceived by the individual. This is a subjective evaluation of the probability that we interpret as the probability of how often the events are occurring together. This is obtained by the conditional probability of two events occur successively in the same probability space. Finally, the H value is normalized by a max-min normalization technique (Eq. 5).

$$SV = H_{norm}({}_{S1}E_{S2}) = \frac{(SL_{S1} +_{S1} CT_{S2} +_{S1} PP_{S2}) - H_{\min}}{H_{\max} - H_{\min}}; \tag{5}$$

The discrepancy (D) can take three distinct values, a positive ($(SV - AV) > 0$), neutral ($(SV - AV) = 0$), and negative ($(SV - AV) < 0$), that are mapped into the following classification: increase, maintain, and decrease, respectively. So, the predictive module of the stress recognition process returns a stress outlook (decrease, maintain, or increase) for the next step of the conflict resolution process. However, our approach should be further developed in order to explore all the formalisms available in CATS.

4 Stress-aware Conflict Resolution

The stress recognition model described previously is being used in the context of ODR. In fact, current research on this field mainly focuses on tools that can facilitate the compilation and exchange of information. This leaves aside very important information. Namely, information about the stress levels or the emotional state is simply not considered and is difficult to be transmitted over the telecommunication means currently used (e.g. chat rooms, e-mail). This not only makes it more difficult for parties to share their ideas but also to perceive feedback from our counterparts [**?**], making communication less effective. Moreover, when parties lack this information, they tend to disregard the fears and desires of each other (as they dont actually see "the person behind the screen"), being more prone to hurt or offend, thus making conflict resolution harder. In that sense, providing feedback about the state of each other is essential in order for parties to maintain a sense of reality and mutual respect. Based on this, we are applying this stress recognition model in our conflict resolution platform. The group of sensors and associated features currently being used includes:

- Touch intensity - higher levels of touch intensity are associated with increased levels of stress;
- Touch accuracy - this is a measure of the amount of touches in active controls versus touches in active areas. When people are nervous or stressed, the accuracy tends to decrease (one of the manifestations of stress are tremors);
- Accelerometer - when people are under stress, they tend to show more abrupt and sudden movements. This can be detected by accelerometers;
- Image processing - the amount of movement of a user can also be analyzed through image processing, allowing to estimate the level of agitation, thus stress;
- Touch patterns - calm touch patterns are generally different from stressed ones. Although touch patterns vary from an individual to another, they are affected by stress in in similar ways. This allows classifying touches as calm or stressed.

In order to have access to this information, we are using portable devices that are equipped with touch screens and sensors and act as interfaces for the conflict resolution platform. Moreover, video cameras placed in front of and around the user allow obtaining additional information. Evidently, not all these sensors may be available at the same time (e.g. some portable devices have capacitive screens, which dont support pressure measurement). In that sense, the afore mentioned techniques are used to select, in each moment, the best set of sensors to use and the weight of each one.

Figure 4 shows the interface of the stress monitor. In this case, only the accelerometer and the touch screen were available. This allows compiling several types of information. Considering the accelerometer, it produces three visualizations of the data: (1) the raw data of the acceleration; (2) the acceleration without the data corresponding to touches and (3) the acceleration during touches. The first one allows seeing all the acceleration information (three axes and absolute value). The second one allows seeing this information without the acceleration due to touches. In fact, variations in acceleration are expected during touches, which will influence the determination of the levels of stress. In that sense, this visualization is more accurate. Finally, it is also possible to examine these variations of acceleration: higher variations in the acceleration during touches are related to increased levels of stress. The interface also shows the touch pattern for the last touch and the corresponding quadratic curve, that allows to compare it with already classified curves and classify the touch as stressed or calm. There is also a screen that shows the evolution of the score. This score is computed based on the performance of the user in doing a given task. In this case, the tasks consist of relatively simple mental calculations. Some stressors are considered to induce stress in the user, specifically the vibration of the device, annoying sounds and a decreasing time to perform the task. All of this influences the performance of the user in completing the tasks, which is reflected on the score, and thus, on the level of stress. The quality of the estimation of the level of stress depends on the amount and quality of the information available at each time.

5 Conclusions

In this work we focused on how to estimate the stress state evolution from the users. This information can then be used by either the platform or even a mediator that is

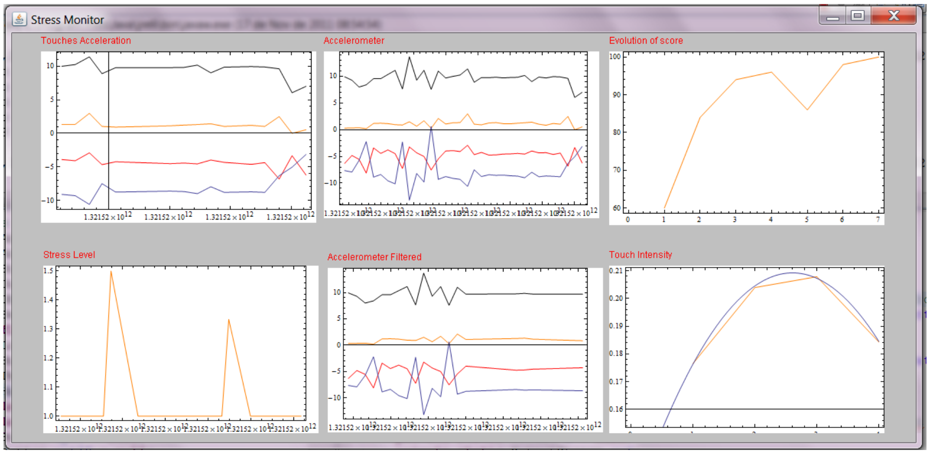

Fig. 4. Prototype of the interface of the Stress Monitor.

conducting process, to perceive how each issue or event is affecting each party. This, we believe, will increase the rate of success of conflict resolution procedures and bring them closer to the rich communicative environment that we have, when we communicate face-to-face. In future work we intend to incorporate additional components to the context and behavioural characterization, so that we can have a more accurate approach. Moreover, in a later phase, we intend to work with the School of Medical Sciences, to use, for instance, electroencephalograms. This will be useful not only for validating this approach but also to more accurately calibrate it.

Acknowledgments. The work described in this paper is included in TIARAC - Telematics and Artificial Intelligence in Alternative Conflict Resolution Project (PTDC/JUR/71354/2006), which is a research project supported by FCT (Science and Technology Foundation), Portugal. The work of Davide Carneiro is also supported by a doctoral grant by FCT (SFRH/BD/64890/2009).